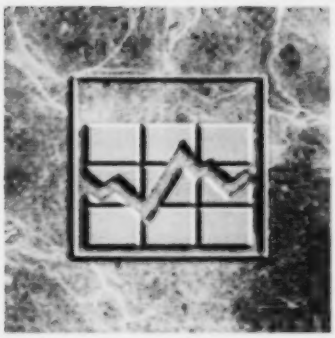
N° 75F002MIF au catalogue — N° 005
ISSN : 1707-2867
ISBN : 978-0-662-08218-7

Document de recherche

Série de documents de recherche - Revenu

Solutions aux problèmes de lisière dans le cadre de l'enquête sur la dynamique du travail et du revenu



par Georges Lemaître

Division de la statistique du revenu
Immeuble Jean-Talon, Ottawa, K1A 0T6

Téléphone : 1-613-951-7355

Statistique Canada Statistics Canada

Canada

SOLUTIONS AUX PROBLÈMES DE LISIÈRE DANS LE CADRE DE L'ENQUÊTE SUR LA DYNAMIQUE DU TRAVAIL ET DU REVENU

Août 1992

Georges Lemaître

(Étude menée sous contrat pour Statistique Canada)

Les Documents de recherche de l'Enquête sur la dynamique du travail et du revenu sont conçus en vue de communiquer les résultats des études ainsi que les décisions importantes prises à l'égard de cette enquête. Ils sont offerts gratuitement, en français et en anglais. Pour obtenir une description sommaire des documents disponibles ou un exemplaire de ces documents, communiquer avec Philip Giles, EDTR, Édifice Jean-Talon, 11ème étage, section D8, Statistique Canada, Ottawa (Ontario), Canada, K1A 0T6.
Téléphone: (613) 951-2891; Télécopieur: (613) 951-3253.

SOMMAIRE

Au cours des dernières années on s'est beaucoup intéressé à ce qui est connu comme le problème "de la lisière" dans les enquêtes de conception longitudinale. Ce problème se rapporte au fait que le nombre de transitions ou de changements dans l'état observé d'un côté à l'autre de la lisière quand on juxtapose les données pour deux périodes de référence consécutives est considérablement plus élevé (parfois un ordre de grandeur plus élevé) que le nombre moyen de transitions ou de changements observés dans les données déclarées pour une seule période de référence.

Les erreurs de réponse sont la cause la plus probable des biais dus à la lisière. Pour des caractéristiques comme la situation vis-à-vis de l'activité ou le fait de toucher un revenu, les erreurs peuvent être dues à des omissions ou au fait de mal situer des événements dans le temps. Cependant, les explications habituelles, données pour les erreurs de réponse, fondées sur la "théorie de l'oubli" ne sont pas corroborées par les données. Les résultats relatifs aux effets de la réponse par procuration sont variables mais ne montrent généralement aucune association claire.

Une analyse conceptuelle des erreurs possibles et des résultats signalés dans les ouvrages publiés sur ce sujet laisse supposer que la tendance des répondants à repousser dans le temps, en particulier au début de la période de référence, des changements qui se sont produits un peu plus tard, peut constituer un élément important dans la production des effets de la lisière.

Pour d'autres caractéristiques comme les salaires, la syndicalisation, les heures habituelles travaillées, les effets de la lisière résulteraient de la variance de réponse ordinaire.

La réalisation d'interviews avec rétro-information (c.-à-d. le fait de rappeler aux répondants les réponses fournies au cours d'une interview antérieure) semblerait la stratégie la plus appropriée pour traiter les effets de la lisière. Toutefois, toutes les techniques de rappel ne donneront pas nécessairement les résultats escomptés. Une comparaison de deux techniques de ce genre, une qui n'a pas réussi à éliminer les effets de la lisière (SIPP) et une qui a connu du succès (EA) et dont l'essai fait dans le cadre de la SIPP a aussi été couronné de succès, tente de déterminer les caractéristiques clés requises.

Dans le présent rapport on soutient, en conclusion, que l'interview avec rétro-information ne devrait pas être considérée comme un mal nécessaire qui est requis pour réduire les biais dus à la lisière, mais comme une partie intégrante du processus d'interview dans les enquêtes longitudinales.

TABLE DES MATIÈRES

		Page
1.	Introduction	1
2.	L'effet de la lisière - Un exemple	2
3.	Une étude des renseignements fournis pour l'effet de la lisière	4
4.	Les effets de la lisière - Pourquoi ils se produisent	6
	4.1 Une analyse conceptuelle des flux de personnes qui deviennent des travailleurs autonomes ou qui cessent de l'être	6
	4.2 La cause probable - Une hypothèse	10
	4.3 Une étude des ouvrages publiés sur le sujet	11
	4.4 Pourquoi ils se produisent - Résumé et conclusion	18
5.	L'interview avec rétro-information - Une solution?	19
	5.1 Une comparaison de deux techniques de rappel	20
	5.2 Autres méthodes qui pourraient être utilisées pour traiter les effets de la lisière	23
6.	Autres résultats provenant de l'enquête sur l'activité	24
7.	Les interviews avec rétro-information - Combien faut-il rappeler de renseignements?	27
	7.1 Quand faut-il rappeler des renseignements et quand faut-il ne pas le faire?	28
	7.2 La rétro-information comme partie intégrante de l'interview	30
8.	Conclusion	31
Bibliographie		33

1. Introduction

Au cours des dernières années on s'est beaucoup intéressé à ce qui est connu comme le problème "de la lisière" dans les enquêtes de conception longitudinale. Ce problème se rapporte au fait que le nombre de transitions ou de changements dans l'état observé d'un côté à l'autre de la lisière quand on juxtapose les données pour deux périodes de référence consécutives est considérablement plus élevé (parfois un ordre de grandeur plus élevé) que le nombre moyen de transitions ou de changements observés dans les données déclarées pour une seule période de référence.

Les taux de transition élevés d'un côté à l'autre de la lisière ne correspondent pas à un phénomène quelconque effectivement observable dans le monde réel et on les a attribués à la présence d'erreurs de déclaration. En particulier, une erreur de déclaration lors d'une de deux occasions distinctes par une personne dont l'état n'a pas changé crée, effectivement, un changement factice. Le nombre de ces changements factices peut être très important par rapport au nombre de changements véritables qui se produisent.

C'est relativement tardivement que les responsables des enquêtes longitudinales ont saisi la signification des erreurs de déclaration et leur incidence sur la mesure des transitions. Il y a déjà un certain temps que le problème est connu parmi les analystes des flux bruts de la population active (voir, par exemple, NCEUS, 1979) et plusieurs méthodes d'estimation ont été proposées pour régler ce problème (PCGF, 1985).

Le présent rapport porte sur le problème de la lisière comme on le rencontre dans le contexte des enquêtes longitudinales. Bien que la structure d'enquête par panel des enquêtes sur la population active se prête à une analyse longitudinale, ces

enquêtes ne sont pas, à proprement parler, de conception longitudinale[1]. Néanmoins, le travail effectué sur les flux bruts a contribué à la compréhension de l'effet des erreurs de déclaration sur la mesure des transitions.

2. L'effet de la lisière - Un exemple

Avant d'étudier le travail déjà effectué dans ce domaine et d'examiner plus attentivement la nature du problème, voyons un exemple.

La figure 1 montre une classification croisée de personnes selon leur état par rapport au travail autonome à deux moments successifs t et t+1, qui peuvent être des mois ou des années. Le nombre de travailleurs autonomes aux moments t et t+1 est, respectivement, a+b et a+c et le changement net dans le nombre de travailleurs autonomes du moment t au moment t+1 est (a+c)-(a+b)=c-b. Ces résultats sont les genres d'estimations produites à partir des enquêtes sur la population active. Avec une enquête longitudinale, parce qu'on dispose de renseignements sur l'état vis-à-vis du travail autonome des répondants à des

Figure 1.

		t+1		
		TA	TNA	
t	TA	a	b	a+b
	TNA	c	d	c+d
		a+c	b+d	

TA: travailleur autonome
TNA: travailleur non autonome

[1] Dans une enquête longitudinale on interviewe les mêmes personnes à plusieurs occasions successives. Les échantillons de l'enquête sur la population active sont généralement des échantillons de logements. Si les résidents d'un logement changent pendant que ce logement fait partie de l'échantillon, les résidents originaux sortent de l'échantillon et sont remplacés par les nouveaux résidents.

moments successifs, on peut aussi obtenir des estimations des cases intérieures a, b, c et d (souvent appelées les "flux" par opposition aux totaux marginaux qui sont appelés les "stocks").

Dans les enquêtes longitudinales, comme on les réalise actuellement, les cases b et c, qui renferment le nombre de personnes qui commencent à être des travailleurs autonomes ou qui cessent de l'être, d'une période à l'autre, tendent à être surestimées d'un côté à l'autre de la lisière (a et d, d'autre part, sont sous-estimées). Il faut remarquer que la variation nette c-b peut encore être non biaisée, si c et b sont surestimées par approximativement la même quantité.

Les estimations réelles observées pour les caractéristiques mentionnées ci-dessus dans l'enquête sur l'activité[2] de Statistique Canada figurent dans les tableaux 1.1 et 1.2, le tableau 1.1 montrant l'estimation du nombre de travailleurs autonomes par mois pour 1988 et 1989 et le tableau 1.2 les mouvements des personnes qui deviennent travailleurs autonomes et qui cessent de l'être pour la même période. La lisière est la période entre le dernier mois visé par l'enquête de 1988 (décembre) et le premier mois visé par l'enquête de 1989 (janvier). Les renseignements pour décembre 1988 ont été recueillis de 1 à 3 mois plus tard, alors qu'une année complète s'était écoulée quand le répondant a fourni les renseignements pour janvier 1989.

Il faut remarquer le sommet dans le nombre de transitions d'un côté à l'autre de la lisière. Cela est typique de la tendance qu'on observe pour des estimations de ce genre. Le nombre de transitions est de trois à quatre fois plus élevé que celui que

---

[2] L'enquête sur l'activité était une enquête longitudinale annuelle réalisée avec deux panels distincts, le premier portant sur les années 1986 et 1987 et le second sur la période allant de 1988 à 1990. Les interviews étaient réalisés de janvier à mars et les renseignements étaient recueillis à propos de l'expérience des répondants sur le marché du travail au cours de l'année civile précédente.

l'on observe pendant le reste de l'année. On observe aussi, en juillet, des mouvements relativement importants de personnes qui cessent d'être des travailleurs autonomes et, en août et septembre, des mouvements importants de personnes qui deviennent des travailleurs autonomes. Cependant, ces mouvements correspondent vraisemblablement à des cessations temporaires réelles d'activité pendant les mois d'été. Sur le marché du travail, il n'existe pas de raison comparable qui expliquerait pourquoi on observe de grandes transitions dans les deux directions de décembre à janvier.

3. Une étude des renseignements fournis pour l'effet de la lisière

C'est Czajka (1983) qui, le premier, a fait remarquer que les transitions tendaient à être plus élevées "entre les cycles" qu'"à l'intérieur d'un cycle"[3] pour ce qui est de la déclaration du fait de toucher un revenu dans le cadre du Income Survey Development Program (ISDP) aux États-Unis, le programme qui a précédé la Survey of Income and Program Participation (SIPP). Moore et Kasprzyk (1984) décrivent des effets de la lisière importants pour une gamme étendue de données élémentaires dans le ISDP, y compris le fait de recevoir des gains au titre des salaires et traitements, des prestations de sécurité sociale, des allocations de chômage, des prestations de pension d'un régime privé, de l'aide aux familles avec enfants à charge (AFEC) et des coupons alimentaires. En fait, les deux seules sources de revenu pour lesquelles l'effet de la lisière n'était pas systématique et important étaient étroitement liées - les prestations de formation et les Basic Educational Opportunity Grants. Les auteurs faisaient l'hypothèse que, puisque ces sources mettaient en jeu des paiements uniques effectués au début des sessions,

---

[3] Le terme "cycle" est utilisé pour décrire une des interviews successives d'une enquête longitudinale. La terminologie "entre les cycles" "à l'intérieur d'un cycle" bien qu'elle soit un peu inélégante, constitue une expression abrégée qui est devenue généralement acceptée dans les publications.

il était plus facile de leur attacher une date que pour les autres sources et que la déclaration exacte du paiement unique ne pouvait jamais produire plus de mouvements entre les cycles qu'à l'intérieur d'un cycle.

Burkhead et Coder (1985) ont démontré des effets semblables dans le cas de la SIPP (avec sa période de référence de quatre mois) pour les gains, les prestations de sécurité sociale, les allocations de chômage, l'AFEC, les prestations de pension d'un régime privé, les coupons alimentaires et Medicaid et cela, en dépit du fait que la SIPP incorpore, dans le questionnaire, une procédure pour mettre à jour l'ensemble des sources de revenu basée sur les données recueillies lors de l'interview du dernier cycle. Hill (1987) signale des effets de la lisière dans la Panel Survey of Income Dynamics (PSID) pour les allocations de chômage et les coupons alimentaires. Des effets de la lisière sont aussi évidents dans l'exemple cité plus tôt sur le travail autonome tiré de l'enquête sur l'activité et pour d'autres caractéristiques comme la recherche d'emploi, le désir de travailler et la formation, que nous examinerons plus loin dans le présent rapport. Finalement, des analyses du changement brut dans la situation vis-à-vis de l'activité laissent supposer que le problème existe même pour des enquêtes avec des périodes de référence aussi courtes qu'un mois.

En résumé donc, il semblerait que les effets de la lisière soient un problème général des enquêtes longitudinales réalisées actuellement, quelles que soient les différences dans leur conception et dans la durée de la période de référence. La prochaine section portera sur les raisons qui sont à la base de l'existence des effets de la lisière.

4. Les effets de la lisière - Pourquoi ils se produisent

Nous avons indiqué plus tôt que les erreurs de réponse sont la cause probable du problème de lisière. Examinons cette situation de plus près. Généralement, on ne s'attendrait pas à ce que les répondants déclarent des événements qui ne se sont jamais produits. Les scénarios beaucoup plus probables sont que les répondants oublient complètement de déclarer certains événements ou les situent mal dans le temps. Il se peut que cette situation s'accompagne ou non d'erreurs dans les mesures de la durée, selon que le fait de mal situer un événement dans le temps augmente ou diminue la durée de l'état qui précède l'événement ou qui lui succède. De plus, des effets de la lisière peuvent se produire dans la mesure de caractéristiques comme les heures habituelles travaillées où la question ne porte pas sur la présence ou l'absence de la caractéristique ou sur le moment où elle est constatée. Dans de tels cas, il est probable que la variance de réponse ordinaire puisse expliquer les effets de la lisière. Nous ne considérerons ces effets que plus tard dans le rapport.

4.1 Une analyse conceptuelle des flux de personnes qui deviennent des travailleurs autonomes ou qui cessent de l'être

Pour que la discussion continue de porter sur le problème qui nous intéresse, retournons à l'exemple déjà mentionné qui portait sur le travail autonome et étudions tout d'abord les flux de personnes qui cessent d'être des travailleurs autonomes à la lisière entre décembre d'une année et janvier de l'année suivante. (Voir Figure 2). Nous étudierons plus tard les flux de personnes qui deviennent des travailleurs autonomes. Il peut se produire, de deux façons, des flux factices, à la lisière, de personnes qui cessent d'être des travailleurs autonomes :

a) si des personnes qui n'étaient pas des travailleurs autonomes en janvier et en décembre déclarent, par erreur, avoir été des travailleurs autonomes en décembre mais déclarent correctement ne pas en être en janvier;

b) si des personnes qui étaient effectivement des travailleurs autonomes en décembre et en janvier déclarent en être en décembre mais ne pas en être en janvier.

Nous supposons ici que l'autre possibilité, nommément celle qu'une véritable transition pour les personnes qui cessent d'être des travailleurs autonomes de décembre à janvier a été inscrite par erreur comme une transition de personnes qui deviennent travailleurs autonomes, est très improbable parce qu'elle exigerait une structure d'erreurs de réponse complexe et peu plausible.

Il faut remarquer qu'ici nous exposons les choses comme si les réponses étaient toujours sans personne interposée. Nous continuerons à exposer les hypothèses de cette façon, bien qu'on pourrait s'attendre à ce qu'une réponse sans personne interposée à une occasion et qu'une réponse par personne interposée lors de l'autre occasion pourraient contribuer à un plus grand nombre de transitions factices. Cependant, des études réalisées au cours des dernières années (Lemaître, 1988; Moore, 1990) ont démontré que l'on a quelque peu exagéré les avantages des réponses sans personne interposée pour ce qui est de la qualité des données. D'autres études mentionnées plus loin portent directement sur l'effet de la déclaration par personne interposée sur le problème de la lisière.

Compte tenu de ce qui a été mentionné précédemment à propos des causes probables des erreurs de déclaration, le cas a) ci-dessus peut se produire si des personnes étaient des travailleurs autonomes plus tôt dans l'année de référence antérieure, ont cessé de l'être avant décembre mais ont déclaré, par erreur, que leur emploi à titre de travailleur autonome s'est prolongé jusqu'en décembre. Cette

situation est traditionnellement décrite par l'expression "télescopage", terme utilisé pour décrire la tendance prétendue des répondants à se rappeler des événements comme s'ils s'étaient produits plus récemment que ce n'est effectivement le cas (Neter et Waksberg, 1965). L'hypothèse du télescopage a été avancée pour expliquer le fait que certaines caractéristiques tendaient à être sous-déclarées au début d'une période de référence et surdéclarées à la fin de cette période. L'événement qui a été télescopé en aval ici est la transition de personnes qui ont cessé d'être des travailleurs autonomes. Le télescopage du flux de personnes qui ont cessé d'être des travailleurs autonome se traduirait par une surdéclaration apparente du nombre de travailleurs autonomes en décembre et peut-être aussi au cours des mois précédents.

Le cas b) peut se produire si des répondants, au moment de la deuxième interview, oubliaient complètement avoir déclaré du travail autonome lors de la dernière interview (ce qui peut se produire si le travail autonome était temporaire ou de nature marginale) ou s'ils ont effectivement cessé d'être des travailleurs autonomes après janvier mais que, de mémoire, leur période de travail autonome remontait à janvier et qu'ils ont déclaré les renseignements en conséquence. Bien entendu, ce dernier cas représente le contraire du télescopage, parce qu'il laisse supposer que des répondants considèrent que des événements se sont produits il y a plus de temps que cela n'est réellement le cas. Nous utiliserons dorénavant l'expression "télescopage inverse" pour désigner cette situation. Analogue au cas a) ci-dessus, le cas b) serait observé comme une sous-déclaration du travail autonome au début de la période de référence.

Par contre, il pourrait se produire des flux factices de personnes qui deviennent des travailleurs autonomes; tout d'abord, si des personnes étaient effectivement des travailleurs autonomes lors des deux mois, mais qu'elles ont déclaré aucun travail autonome en décembre parce que, de mémoire, la transition au travail autonome

s'était produite plus tard qu'en décembre. Cela semble probable seulement si les interviews sont réalisées un certain temps après la fin de la période de référence (pour la PSID, par exemple, les interviews sont réalisées de mai à septembre inclusivement). Le deuxième cas est celui des personnes qui ne sont des travailleurs autonomes, ni en décembre, ni en janvier et qui déclarent, par erreur, du travail autonome en janvier. Cela pourrait se produire si la transition au travail autonome s'est produite plus tard dans l'année, mais qu'elle est déclarée comme si elle s'était produite en janvier, au début de la période de référence. Il s'agit, à nouveau, de télescopage inverse.

Pour résumer donc, il peut y avoir des transitions erronées de personnes qui cessent d'être des travailleurs autonomes si les répondants :

i) déclarent, avec raison, aucun travail autonome en janvier mais, lorsqu'ils déclarent les données pour la première période de référence,
   - télescopent jusqu'en décembre du travail autonome qui a pris fin plus tôt au cours de la période de référence antérieure;

ii) déclarent, avec raison, du travail autonome en décembre mais, lorsqu'ils déclarent les données pour la deuxième période de référence,
   - oublient complètement le travail autonome déclaré lors du premier cycle;
   - font du télescopage inverse jusqu'en janvier, pour une transition hors du travail autonome plus tard dans la deuxième période de référence.

Par contre, il y aura de fausses transitions au travail autonome, si les enquêtés :

i) déclarent des données exactes pour janvier mais, lorsqu'ils déclarent les données pour la première période de référence,
   - télescopent une transition au travail autonome, qui s'est effectivement produite pendant la première période de référence,

jusqu'à l'année suivante (cela ne s'applique que lorsque les interviews sont réalisées après décembre);

ii) déclarent des données exactes pour décembre mais, lorsqu'ils déclarent les données pour la deuxième période de référence,

- font du télescopage inverse jusqu'en janvier, pour une transition au travail autonome qui s'est produite plus tard dans la deuxième période de référence.

Il faut remarquer que pour les flux de personnes qui cessent d'être des travailleurs autonomes, il existe une source d'erreur (le fait d'oublier une période de travail autonome) qui n'a pas d'équivalent dans les sources d'erreur pour les flux de personnes qui commencent à être des travailleurs autonomes. De plus, dans ce dernier cas, la source i) semble être peu probable quand les interviews sont réalisées immédiatement après la fin de la période de référence. D'après ces observations, on pourrait prédire plus de faux négatifs (c.-à-d. le fait de ne pas déclarer une période de travail autonome au cours d'un mois donné) que de faux positifs (le fait de déclarer du travail autonome pour un mois donné quand il n'y en a pas eu). Comme nous le verrons plus loin, pour le fait de toucher un revenu, ce sera exactement ce que Marquis, Moore et Huggins (1990) ont observé pour la SIPP.

4.2 La cause probable - Une hypothèse

Dans les enquêtes longitudinales, la période dont le répondant doit se souvenir est plus courte (dans les enquêtes annuelles, considérablement plus courte) pour les données recueillies à propos du mois précédant la lisière que pour les données du mois qui suit la lisière. Par conséquent, si, en fait, des effets liés aux erreurs de mémoire entrent en jeu, on s'attendrait à ce qu'il y ait plus d'erreurs de déclaration pour les données de janvier, dans notre exemple, que pour les données de

décembre. Cela laisserait supposer que le télescopage inverse, vers la lisière, des transitions qui se sont produites plus tard dans la deuxième période de référence pourrait être le principal facteur à l'origine des effets de la lisière. Pour les flux de personnes qui cessent d'être des travailleurs autonomes, il y a la possibilité additionnelle que, lors de la deuxième interview, les répondants oublient complètement la période de travail autonome qui a eu lieu et qui a été déclarée lors de l'interview antérieure.

Il existe aussi d'autres facteurs qui pourraient contribuer aux effets de la lisière. Il est possible que le travail autonome ait été inscrit inexactement en premier lieu, c'est-à-dire, qu'il n'y en ait jamais réellement eu. Toutefois, comme nous l'avons déjà mentionné, on ne s'attendrait pas normalement à ce que des répondants déclarent des incidents qui n'ont jamais eu lieu. D'autres possibilités sont de nature plus aléatoire et comprennent, par exemple, une mauvaise compréhension de la question par le répondant ou de la réponse par l'intervieweur, une déclaration intentionnellement inexacte par le répondant et des erreurs de codage, de traitement ou d'imputation.

Avec ces renseignements généraux, il est intéressant d'examiner ce que la recherche dans le domaine des origines et des causes des effets de la lisière a découvert.

4.3 Une étude des ouvrages publiés sur le sujet

Moore et Kasprzyk (1984) examinent la possibilité que les effets de la lisière puissent refléter des tendances sous-jacentes réelles, des erreurs de contrôle et d'imputation ou des erreurs dans l'appariement des enregistrements provenant de deux interviews. Toutefois, ils écartent la possibilité qu'une quelconque de ces situations contribue sensiblement au problème de la lisière et ils reprennent

l'explication la plus courante donnée pour ces effets, nommément une certaine forme de biais causé par les erreurs de mémoire.

Selon eux, si les répondants ont tendance, rétrospectivement, à déclarer moins de changements qu'il ne s'en est réellement produit pendant une période de référence, alors cela aurait tendance à produire, de deux façons, plus de mouvements entre les cycles qu'à l'intérieur d'un cycle, tout d'abord en réduisant l'incidence des transitions à l'intérieur d'un cycle et deuxièmement, en déplaçant vers la lisière entre les cycles l'occurrence des transitions. On cite une étude de Goudreau, Oberheu et Vaughan (1984) qui donne un certain appui à ces notions. Cette étude de bénéficiaires connus de l'AFEC a montré que a) les personnes qui n'ont pas déclaré avoir reçu des prestations d'AFEC ne les avaient vraisemblablement reçues que pendant une partie de la période de référence et que b) l'erreur la plus courante pour la déclaration de prestations était la tendance à déclarer le montant le plus récent pour tous les [quatre] mois de la période de référence quand les sommes avaient, en fait, varié. Dans les deux cas, la tendance des répondants est de déclarer moins de variation pendant la période de référence qu'il n'y en a réellement eu.

Une deuxième explication du problème de la lisière examiné par Moore et Kasprzyk porte sur l'effet de la déclaration par personne interposée. Ces derniers ont formulé l'hypothèse qu'il était plus probable qu'il se produise des changements factices dans l'état quand des répondants différents ont fourni les données lors des deux occasions. Toutefois, les auteurs n'ont pas trouvé de preuve que l'autodéclaration lors de cycles consécutifs éliminait les effets de la lisière. En fait, les preuves disponibles laissent supposer que l'autodéclaration lors des deux occasions a empiré l'effet de la lisière!

Dans une étude réalisée plus tard, Weidman (1986) a étudié la possibilité que des covariables démographiques (âge, sexe, scolarité, état matrimonial, etc.) ou l'état de l'interview (combinaisons de déclarations par personne interposée ou sans personne interposée) pourraient aider à expliquer les mouvements plus élevés entre les cycles observés dans la SIPP, mais il n'a trouvé aucune preuve concluante d'une association quelconque.

Hill (1987) étudie l'effet de la lisière pour la PSID dans le contexte d'une année (1984) pour laquelle il y a eu chevauchement de la lisière dans le plan d'enquête. C'est-à-dire que les données pour janvier 1984 ont été recueillies à la fois par l'enquête qui portait normalement seulement sur l'année civile 1983 et par celle qui portait sur l'année civile 1984. Ainsi, on peut examiner les transitions de décembre 1983 à janvier 1984 telles qu'elles ont été déclarées à une occasion ou lors des deux occasions. L'analyse portait sur la situation vis-à-vis de l'activité. Les déclarations des enquêtés relativement à la situation vis-à-vis de l'activité ont été classées en trois groupes selon la cohérence entre le rapport de 1984 pour janvier 1984 et celui qui a été fourni en 1985, de la façon suivante :

a) rapports cohérents;

b) incohérences alors que le rapport de 1985 n'indiquait aucun changement dans la situation depuis décembre 1983;

c) incohérences où le rapport de 1985 montrait un changement de situation.

Il faut remarquer que le cas c) tend à contribuer à un effet de la lisière, alors que le cas b) réduit l'incidence d'un tel effet.

Une analyse des facteurs qui pouvaient déterminer l'incohérence laissait supposer que les Noirs, les répondants plus âgés et les personnes avec des durées d'emploi plus courtes tendent à contribuer à l'effet de la lisière, alors que les femmes et les personnes avec des revenus plus élevés tendent à répondre de façon à atténuer cet effet. Par contre, la durée de la période dont le répondant doit se souvenir (les

interviews pour la PSID sont réalisées sur une période de six mois) et le nombre de transitions qui se sont produites au cours de l'année de référence 1984, deux facteurs qui sont normalement avancés comme des mesures de la difficulté de déclaration auxquelles le répondant doit faire face, n'ont eu aucun effet perceptible.

Toutefois, l'auteur admet avoir été incapable de fournir une explication causale adéquate des résultats observés. Dans les analyses de ce genre, où l'on utilise de nombreuses variables qui pourraient être «explicatives» et entre certaines desquelles il pourrait y avoir intercorrélation, les effets peuvent être répartis sur un certain nombre de variables, de sorte qu'il est difficile de déterminer quel est le facteur significatif sous-jacent (s'il y en a un). Une étude (Lemaître, 1988) a montré que les erreurs de réponse relatives à la situation vis-à-vis de l'activité tendent à être associées à des heures moins longues et à des durées d'emploi plus courtes (c.-à-d. à un emploi moins stable). Il est raisonnable de supposer que les répondants qui ont des antécédents d'emploi instable auront plus de difficulté à déclarer leur emploi avec exactitude. Toutefois, dans les analyses de régression, il se peut que des covariables de l'emploi instable (telles que le faible niveau de scolarité, la race, le faible revenu) puissent capter certains des effets. Il se peut que ce soit ce qui est reflété dans l'analyse de Hill.

Kalton et Miller (1991) étudient la déclaration des montants de sécurité sociale dans la SIPP lors d'une période (janvier 1984) où il y a eu une augmentation réelle (de 3.5 %) des prestations. L'échantillon de la SIPP est composé de quatre groupes de renouvellement qui sont interviewés lors de mois consécutifs, les quatre mois antérieurs constituant la période de référence. Ainsi, les données pour des paires consécutives de mois sont recueillies dans le même cycle pour trois des quatre groupes de renouvellement mais dans des cycles différents pour le quatrième. Cela permet d'effectuer une comparaison des déclarations à l'intérieur d'un cycle et entre les cycles pour toute paire de mois consécutifs.

Les auteurs signalent l'excès habituel de transitions entre les cycles par rapport aux transitions à l'intérieur d'un cycle. Ils étudient ensuite de plus près la déclaration de l'augmentation, en janvier 1984, des prestations pour les trois groupes de renouvellement dont les périodes de référence comprenaient à la fois décembre 1983 et janvier 1984. Ils trouvent, tout d'abord, une association entre la durée de la période de remémoration et la proportion de répondants qui ont déclaré aucun changement dans les prestations; c'est-à-dire que, plus l'augmentation s'était produite longtemps avant la date de l'interview, moins il était probable qu'elle soit déclarée.

Ils modélisent aussi la déclaration exacte de cette augmentation de 3.5 % (définie comme étant toute déclaration entre 2 % et 4.1 %) comme fonction d'un ensemble de variables explicatives et trouvent une association entre la déclaration exacte et la récence du mois de référence, l'autodéclaration, la race et le fait de recevoir des prestations plus élevées de sécurité sociale. Une question qui se présente relativement à la non-déclaration de l'augmentation des prestations est le fait de savoir si l'augmentation aurait pu être déclarée à un autre moment, en particulier lors de la lisière précédente ou suivante. Kalton et Miller ont trouvé une tendance significative pour que cela se produise, particulièrement à la lisière précédente. En d'autres mots, un nombre significatif de répondants déclarent que le montant qu'ils reçoivent actuellement est celui qu'ils ont reçu au cours de toute la période de référence. Cela laisse supposer qu'une partie du problème de la lisière est due au fait que les répondants ont tendance à faire du télescopage inverse avec la date où les événements se produisent, c'est-à-dire qu'ils perçoivent les événements comme s'ils s'étaient produits plus tôt que ce n'est effectivement le cas et ce, surtout au début de la période de référence visée par l'enquête.

Les analyses les plus détaillées du problème de la lisière sont probablement celles de Marquis et Moore (1989) et de Marquis, Moore et Huggins (1990) sur des

données de la SIPP. Ces auteurs ont examiné la déclaration pour huit programmes dont la participation était déclarée par les ménages participant à la SIPP dans quatre états. On a établi un micro-couplage entre les déclarations pour l'enquête et les données administratives pour les mêmes programmes, ce qui a permis d'effectuer une vraie validation des réponses à l'enquête. Les résultats montrent des problèmes de réponse pour pratiquement tous les programmes. En particulier, on a relevé une sous-déclaration de la participation pour six des huit programmes, avec des biais allant de 8 % à près de 40 %. De plus, les transitions étaient presque toujours trop faibles quand elles n'étaient pas déclarées à la lisière et trop fortes quand elles l'étaient, par rapport à ce qui était observé dans les dossiers administratifs.

Les auteurs ont ensuite étudié des hypothèses, provenant de la théorie de l'oubli, qui sont traditionnellement avancées pour expliquer les genres de tendances observées. Selon ces hypothèses, les erreurs sont surtout des omissions ou des sous-déclarations; les sous-déclarations empirent à mesure que le temps écoulé augmente et le rappel des événements récents est exact. Cependant, les résultats de la vérification des dossiers ne corroborent pas ces affirmations. Bien que la sous-déclaration de la participation prédominait, il y a aussi eu beaucoup de surdéclaration pour tous les programmes. De plus, à l'exception des allocations d'assurance-chômage, il n'y avait pas d'amélioration importante dans la déclaration pour le mois le plus récent par rapport à la déclaration quatre mois auparavant.

Voici certains des autres résultats de leur étude :

- Il y a confusion à propos du nom des programmes, mais cette confusion a une importance mineure.
- Il y a plus de sous-déclaration que de surdéclaration à la fois chez les répondants et ceux qui répondent par personne interposée. La sous-

déclaration est plus importante par personne interposée que pour les répondants comme tels.

- Les effets dus à l'intervieweur ne sont pas plus importants que dans d'autres enquêtes.

Malheureusement, les auteurs n'ont pas considéré en détail la possibilité que la surdéclaration observée au cours d'un mois particulier puisse refléter une erreur de classement d'un revenu de la part des répondants qui ont effectivement reçu des sommes au cours d'autres mois ou le prolongement de la réception d'un tel revenu. Bien qu'ils examinent l'importance du télescopage externe (la possibilité que des sommes reçues au cours de la période de référence antérieure soient reportées à la période courante), ils étudient cette situation dans un cadre plutôt limité. Ils s'attendaient à ce que, puisque au cours des cycles suivants, l'intervieweur rappelle au répondant ce que ce dernier a déclaré lors du cycle précédent, il soit peu probable que le répondant déclarerait pour le cycle courant, une participation qui s'est effectivement produite au cours d'un cycle antérieur. La seule exception à cette situation était pour le premier cycle, quand rien n'était rappelé au répondant. Les auteurs s'attendaient donc à trouver une plus grande surdéclaration de la participation lors du premier cycle, mais ils n'en ont trouvé aucune preuve.

Toutefois, le fait de rappeler au répondant ce qui a été déclaré au cours d'un cycle antérieur empêcherait ce dernier de prolonger, par erreur, la participation à un programme dans au moins les premiers mois du cycle courant.

4.4 Pourquoi ils se produisent - Résumé et conclusion

En résumé donc, les études qui ont tenté de déterminer l'origine des effets de la lisière n'ont pas réussi à trouver une cause sous-jacente claire pour les résultats observés. Les résultats relatifs aux effets des déclarations par personne interposée sont contradictoires mais généralement négatifs. Les hypothèses relatives aux effets liés aux erreurs de mémoire n'ont généralement pas été corroborées par les données. Par contre, deux études (Goudreau, Oberheu et Vaughan, 1984 et Kalton et Miller, 1991) laissent supposer que les répondants ont tendance à sous-déclarer des changements au cours d'une période de référence et, en fait, à faire du télescopage inverse jusqu'au début de la période de référence pour le changement qui s'est produit. Cela est exactement le phénomène relevé précédemment comme étant vraisemblablement le principal élément qui contribue au problème de la lisière.

Il est intéressant de constater qu'un télescopage inverse jusqu'à la lisière des flux de personnes qui cessent d'être des travailleurs autonomes montrerait, toutes autres choses étant égales par ailleurs, une sous-déclaration apparente du travail autonome au début de la période de référence. Toutefois, cela serait compensé par le phénomène contraire, nommément un télescopage inverse, jusqu'au mois de la lisière, des flux de personnes qui deviennent des travailleurs autonomes, ce qui tendrait à montrer une surdéclaration du travail autonome au début de l'année. Dans l'ensemble, aucun ou très peu d'effet lié aux erreurs de mémoire ne serait apparent. La présence à la fois de surdéclaration et de sous-déclaration ainsi que l'absence apparente d'effets liés aux erreurs de mémoire, dans l'ensemble, sont précisément ce que Marquis, Moore et Huggins (1990) ont observé dans leur étude qui a fait appel à la vérification des dossiers.

Une autre conséquence du télescopage inverse du changement (soit des personnes qui deviennent des travailleurs autonomes, soit des personnes qui cessent de l'être) jusqu'au mois de la lisière serait une sous-déclaration globale du changement pendant la période de référence, résultat confirmé, à nouveau, par l'étude portant sur la vérification des dossiers de la SIPP. De plus, s'il est plus probable qu'il se produise un télescopage inverse des changements dans le cas de changements moins récents, il y aurait plus de sous-déclaration du changement au début de la période de référence, résultat observé par Kalton et Miller (1991).

Si, effectivement, la caractéristique essentielle de l'effet de la lisière est la déclaration à la lisière d'un changement qui s'est effectivement produit à un autre moment, alors il semblerait qu'un certain genre de rappel aux répondants des réponses données lors de l'interview antérieure (c.-à-d. l'interview avec rétro-information) soit justifié. La prochaine section porte sur cette question.

5. L'interview avec rétro-information - Une solution?

On a rarement tenté de réaliser des interviews avec rétro-information à cause des craintes qu'une telle méthode tendrait à biaiser les estimations du changement. Il faut toutefois évaluer ces craintes dans un contexte approprié. On pourrait bien être peu disposé à effectuer des interviews avec rétro-information dans une situation où l'on sait que les estimations du changement sont non biaisées. Mais dans le cas présent, les données qui nous intéressent sont de nature longitudinale et les estimations des transitions d'un côté à l'autre de la lisière sont fortement biaisées à la hausse. L'incidence prévue de l'interview avec rétro-information va précisément dans la direction souhaitée. Que l'interview avec rétro-information réduise excessivement ou non le changement brut et que cette méthode induise ou non des biais dans les estimations du changement net sont des questions pour

lesquelles nous ne possédons pas encore de réponses claires. Nous reprendrons ces points plus loin.

Remarquons, tout d'abord, que l'interview avec rétro-information peut être réalisée de différentes façons, dont il se peut que certaines ne soient pas couronnées de succès ou ne produisent pas les résultats désirés. Nous mettrons en opposition ici deux techniques de ce genre, une qui a été utilisée, sans succès, dans le cadre de la SIPP depuis 1983 et l'autre qui a été mise en application avec succès dans le cadre de l'enquête sur l'activité en 1987.

5.1 Une comparaison de deux techniques de rétro-information

Dans la SIPP, l'intervieweur lit une liste des sources du revenu reçu pendant au moins un mois de la dernière période de référence. On vérifie l'exactitude de la liste et l'on corrige tout écart. On demande ensuite au répondant s'il a reçu, à un moment quelconque de la période de référence courante, un revenu provenant d'une quelconque des sources de revenu mentionnées pour la dernière période de référence. Finalement, on détermine les sommes reçues pour des mois particuliers quand les montants sont inscrits, à compter du mois le plus récent (Burkhead et Coder 1985). Comme on l'a mentionné plus haut, cette forme d'interview avec rétro-information a eu remarquablement peu de succès pour assurer des transitions régulières à la lisière.

Par contre, dans le cadre de l'enquête sur l'activité (EA), on a recueilli des renseignements sur tous les emplois occupés pendant l'année de référence, y compris les emplois occupés jusqu'au moment de l'interview, qui était réalisée au début de l'année suivante. On agissait de la sorte afin de relever les emplois inscrits au début de l'année suivante et ce sont ces renseignements qui sont rappelés aux répondants lors de l'interview réalisée l'année suivante. On demande

alors aux répondants de confirmer ces renseignements qui, à ce moment, sont utilisés comme point de départ de l'interview pour l'année courante. Une certaine proportion de répondants ne confirment pas les emplois déclarés auparavant et, pour l'EA, on a élaboré des procédures d'imputation qui prolongent les emplois dans l'année courante d'une façon compatible avec l'état de ces emplois tel qu'enregistré au moment de l'interview de l'année précédente. Cette façon d'agir a pour but d'éviter une confrontation avec le répondant, tout en attribuant une pondération plus élevée, lors du traitement, aux renseignements recueillis pendant l'interview de l'année précédente, quand la période de remémoration était beaucoup plus courte. Les estimations résultantes des flux de personnes qui cessent d'être employées à la lisière ont montré une absence complète des effets de la lisière habituels (Murray, Michaud et al., 1991). De plus, les estimations des flux de personnes qui cessent d'être employées ont montré beaucoup moins de sous-déclaration au début de l'année par rapport à d'autres sources que lors du premier cycle de l'EA, pour lequel il n'y avait pas eu de rappel d'information.

La caractéristique clé de la technique utilisée pour l'EA est que les renseignements rappelés aux répondants surmontent effectivement la lisière, parce qu'il s'agit de renseignements pour l'année de référence courante. Ainsi, les répondants, afin de déclarer un changement à la lisière, doivent essentiellement nier un rapport fourni une année auparavant dans des conditions de rappel beaucoup plus favorables. Néanmoins, certains répondants le font (environ 2 %), mais les procédures d'imputation semblent très bien traiter ces cas. Il faut remarquer que l'on n'a pas trouvé que le fait de nier des emplois déclarés auparavant était associé à la réponse par personne interposée dans l'EA. Des tests récents de cette technique, réalisés dans le cadre de la SIPP pour le fait de toucher un revenu, n'ont pas non plus montré d'effet de la lisière (Marquis, 1992).

La technique d'interview avec rétro-information utilisée pour la SIPP, par contre, rappelle aux répondants ce qu'ils ont déclaré pour la période de référence antérieure, avant de leur poser des questions sur les sommes qu'ils ont reçues à un moment quelconque de la période de référence courante. À première vue, cela peut sembler une technique utilisable. Toutefois, cette technique n'empêche pas les répondants de faire du télescopage inverse jusqu'à la lisière pour tout changement qui peut s'être produit pendant la période de référence courante. De plus, on rappelle aux répondants les sources des revenus qu'ils ont reçus pendant tous les mois de la période de référence précédente, pas seulement pendant le dernier mois. Ainsi, il se peut que les répondants qui n'avaient pas reçu une somme d'une source donnée lors du dernier mois de la période de référence antérieure prolongent, par erreur, la durée de la réception des sommes provenant de cette source jusqu'à la période courante. Il est possible que ces personnes déclarent alors avoir reçu des sommes pendant le premier mois de la période, contribuant ainsi à un effet de la lisière.

Les résultats des deux enquêtes sont compatibles avec l'hypothèse que le télescopage inverse du changement jusqu'à la lisière peut bien être la cause principale du problème de la lisière. Si c'est effectivement le cas, alors il est probable que la réalisation d'interviews avec rétro-information aura une incidence très favorable sur les estimations du flux à la lisière (comme cela a été démontré pour l'EA et dans des essais récents portant sur la SIPP) sans avoir d'effet défavorable sur les estimations du changement net ou sans atténuer excessivement le changement brut tant à la lisière que par la suite. Cela est dû au fait que le télescopage inverse n'implique pas qu'il y a parmi les répondants une tendance générale à sous-déclarer un changement, situation que l'interview avec rétro-information ne ferait qu'empirer, mais simplement qu'ils situent mal ce changement dans le temps. Une fois qu'on leur rappelle leur situation au début de la période de référence courante (telle qu'elle a été déclarée lors de l'interview précédente), on

s'attendrait à ce que la situation qui aurait normalement été déclarée à ce moment, si aucun renseignement n'avait été rappelé au répondant soit normalement déclarée un peu plus tard au cours de la période de référence.

5.2 Autres méthodes qui pourraient être utilisées pour traiter les effets de la lisière

En plus de l'interview avec rétro-information, on a aussi proposé des techniques d'estimation pour traiter le problème de la lisière. Toutefois, ces techniques font appel à des renseignements auxiliaires et on ne peut généralement les appliquer qu'une variable à la fois, ce qui ne les rend pas pratiques pour une utilisation normale. Certaines de ces techniques sont mentionnées dans Marquis, Moore et Huggins (1990). Une étude mentionnée par les mêmes auteurs, au cours de laquelle des employés du Census Bureau ont, en 1989, accompagné les intervieweurs de la SIPP quand ces derniers réalisaient leurs interviews présente plus d'intérêt. À des moments appropriés, les employés du Census Bureau interrompaient les interviews pour poser des questions aux répondants à propos du processus de réponse, résumant par la suite les interactions verbales importantes qu'ils observaient.

Une des principales conclusions des résumés est que de nombreux enquêtés adoptent une règle heuristique simple pour répondre rapidement à des questions portant sur des événements qui se reproduisent régulièrement pendant la période de référence de quatre mois de la SIPP. Les répondants utilisent cette règle simple plutôt que de faire un effort de mémoire intense et direct et au lieu de consulter leurs dossiers personnels. Les auteurs concluent que la qualité des réponses pourrait être améliorée si l'on «enseignait» aux répondants comment fournir de meilleurs renseignements (y compris l'emploi de dossiers), si l'on insistait auprès des intervieweurs sur l'importance de la qualité des réponses et si l'on s'assurait que

les intervieweurs expliquent aux répondants les buts de chaque section du questionnaire. Bien qu'il soit fort possible que ces mesures améliorent la qualité des réponses et qu'elles puissent réduire les genres d'erreurs qui mènent à des effets de la lisière, elles ont des implications plutôt importantes pour ce qui est des coûts et elles peuvent ne pas être tout à fait pratiques pour cette raison.

6. Autres résultats provenant de l'enquête sur l'activité

Nous avons déjà mentionné la surestimation à la lisière dans l'EA pour les mouvements de personnes qui deviennent des travailleurs autonomes ou qui cessent de l'être ainsi que l'absence d'effet de la lisière pour les flux de personnes qui cessent d'être employées quand des renseignements sur les emplois sont rappelés aux répondants d'une façon particulière. Dans la présente section, nous étudions d'autres exemples d'effets de la lisière provenant de cette enquête, nous démontrons la nécessité d'effectuer une «pré-interview» pour recueillir des renseignements de base à rappeler aux répondants lors du premier cycle et nous examinons des caractéristiques pour lesquelles les effets de la lisière ne sont pas tellement évidents.

Les graphiques 2.1, 2.2, 3.1, et 3.2 montrent des estimations mensuelles des stocks et des flux des recherches d'emploi et du désir de travailler parmi les personnes sans emploi pendant toute l'année. La question à propos du désir de travailler n'a été posée qu'aux personnes qui ont déclaré n'avoir fait absolument aucune recherche d'emploi pendant l'année de référence. Dans les deux cas, les estimations des flux montrent les effets de la lisière habituels. Toutefois, bien que la forme de la courbe des estimations des stocks pour 1988 soit à peu près comparable à celle de la même courbe pour 1989 pour les deux caractéristiques, ces deux courbes montrent aussi une chute soudaine dans les estimations des stocks de décembre 1988 à janvier 1989. Même si la période dont les enquêtés

doivent se rappeler pour l'estimation de décembre est beaucoup plus courte que pour l'estimation de janvier, la diminution ne peut découler d'une différence dans la façon dont les personnes se rappellent les événements au cours de ces deux mois, parce que toutes les courbes pour 1989 ont été déplacées à la baisse. Qu'est-ce qui pourrait causer ce déplacement?

La recherche d'emploi et le désir de travailler inscrits ici s'appliquent à des personnes qui n'ont déclaré aucun travail durant l'année. L'explication la plus plausible du déplacement est qu'il y a eu une diminution dans le nombre de personnes inscrites, en 1989, comme sans travail pendant toute l'année. Maintenant, des renseignements sur les emplois inscrits en janvier 1989 ont été rappelés aux répondants pour l'interview de 1989. Aucune procédure de ce genre ne s'appliquait pour l'interview de 1988, qui était le premier cycle de l'enquête. Il est vraisemblable que des emplois (probablement marginaux) que les répondants auraient autrement omis de déclarer pour 1989 (soit parce qu'ils ont oublié les avoir jamais occupés, soit parce que, par erreur, ils se rappellent ne plus les avoir occupés en janvier) ont été effectivement déclarés parce qu'on a rappelé aux répondants que ces derniers les occupaient au début de l'année. Ainsi, le fait de rappeler des réponses sur les emplois a réduit le nombre de personnes inscrites comme ayant occupé absolument aucun emploi en 1989 par rapport à 1988.

Les variations dans les stocks nous permettent aussi de déterminer une limite inférieure grossière pour le nombre d'emplois additionnels déclarés en 1989 à la suite de l'utilisation de l'interview avec rétro-information. Cette augmentation est approximativement égale à la différence entre la somme des estimations observées des stocks pour les deux caractéristiques de décembre à janvier et la somme du véritable changement net dans les deux caractéristiques pendant la même période. Les estimations sont additionnées pour les deux caractéristiques parce qu'elles s'appliquent à des groupes de répondants différents; c'est-à-dire qu'on a posé une

question à des personnes à propos de leur désir de travailler seulement si ces dernières avaient précisé n'avoir fait absolument aucune recherche d'emploi pendant l'année. Nous ne savons pas réellement quel était le changement net véritable de décembre à janvier mais, pour obtenir une estimation grossière, nous pouvons attribuer le plus gros changement net qui s'est produit pendant la période de 1988 à 1989 (à l'exclusion, bien entendu, du changement de décembre à janvier) à décembre-janvier. L'estimation du nombre d'emplois additionnels s'élève à un peu moins de 100,000 emplois. Il faut remarquer que ce chiffre est aussi le nombre d'emplois que ce groupe aurait contribué sous forme de flux factices de personnes qui cessent d'être employées si leurs emplois n'avaient pas été relevés en 1989.

Ces exemples particuliers montrent la nécessité de rappeler des renseignements sur le premier cycle, ainsi que sur les cycles successifs, si l'on veut éviter des variations dans les stocks du genre de ceux qui ont été relevés plus haut pour l'EA.

Les figures 4.1, 4.2, 5.1, et 5.2 montrent les estimations des stocks et des flux pendant la période 1988-1989 pour les personnes qui fréquentent l'école à plein temps ainsi que pour les personnes qui suivent des cours de formation professionnelle à plein temps ou à temps partiel. Pour les données sur la fréquentation scolaire, on relève un petit pic à la lisière, mais ce phénomène est évidemment éclipsé par les flux réels de personnes qui commencent ou qui cessent de fréquenter l'école à plein temps qu'on observe pendant l'année. De plus, bien que les flux à la lisière soient considérablement plus élevés que ceux pour les mois adjacents, une bonne partie de ces flux peut, en fait, être réelle, parce que décembre et janvier sont des mois de transition naturels pour la fréquentation scolaire. Dans des cas comme ceux-là, où il y a des aides-mémoire institutionnalisés évidents qui aident le répondant à situer les événements dans le temps, il se peut que le nombre de flux factices soit, en fait, très petit.

De même, on relève des petits pics à la lisière pour les flux de personnes qui suivent des cours de formation professionnelle, mais il n'est pas immédiatement évident si ces flux sont factices ou non, parce qu'il existe des flux aussi considérables à d'autres périodes de l'année. Toutefois, les données fournissent certains indices. On s'attendrait à ce que les flux considérables de personnes qui commencent à suivre des cours de formation professionnelle observés d'août à septembre soient suivis, un peu plus tard, par des flux considérables de personnes qui cessent de suivre des cours de formation quand la formation prend fin. La durée des périodes de formation peut être variable, ce qui signifierait des sorties échelonnées et aucun sommet très évident. Ou un bon nombre de périodes de formation peuvent se dérouler simultanément, avec un sommet soudain dans les sorties. Dans notre cas, on observe un sommet dans les sorties qui se produit exactement à la lisière, ce qui laisse supposer une période de formation allant de septembre à décembre. De même, le sommet dans les entrées de personnes pour suivre des cours de formation professionnelle observé à la lisière n'est pas incompatible avec les sorties de personnes qui cessent de suivre de tels cours qu'on observe pendant les mois suivants.

Néanmoins, l'hypothèse selon laquelle les flux observés à la lisière sont réels comporte des risques et l'on aurait tort de renoncer à l'interview avec rétro-information dans ce cas sans renseignements indépendants à l'appui.

7. Les interviews avec rétro-information - Combien faut-il rappeler de renseignements?

Les exemples portant sur la fréquentation scolaire et sur la formation professionnelle mentionnés ci-dessus nous amènent à nous demander dans quelle mesure on doit avoir recours à l'interview avec rétro-information dans une enquête longitudinale afin de nous assurer d'obtenir des estimations fiables. Doit-on

rappeler les renseignements portant sur toutes les caractéristiques d'un répondant comme ils ont été observés lors d'interviews réalisées auparavant?

7.1 Quand faut-il rappeler des renseignements et quand faut-il ne pas le faire?

La fréquentation scolaire est un exemple de caractéristique où il peut ne pas être nécessaire d'avoir recours à la rétro-information, mais on a beaucoup de difficulté à trouver d'autres caractéristiques où un cadre «naturel» comparable existe pour situer les événements dans le temps. Le cas de la formation professionnelle, par exemple, est moins évident.

On peut probablement supposer que certaines caractéristiques (p. ex., si un répondant occupe ou non un emploi syndiqué ou s'il participe à un régime de pension) ne varient pas à moins que le travailleur change d'employeur ou qu'il soit promu cadre. Ainsi, il n'est pas nécessaire de poser de telles questions dans les interviews réalisées après la première, à moins qu'il n'y ait une modification d'état qui pourrait entraîner un changement dans la caractéristique. En fait, on court après les difficultés quand on pose, à nouveau, de telles questions sans avoir recours à l'interview avec rétro-information, particulièrement si aucun changement dans l'état ne justifie le fait de poser ces questions à nouveau. Il faut remarquer qu'ici, comme dans les autres exemples cités plus loin, les effets de la lisière ne découleraient vraisemblablement pas de l'oubli ou du fait de mal situer des événements dans le temps, mais de la variance de réponse ordinaire.

Les totalisations du changement dans les salaires pour les salaires reçus au même emploi mais déclarés indépendamment lors de deux interviews réalisées à une année d'intervalle laissent supposer qu'il serait recommandable de rappeler des renseignements sur les salaires (Bootby, 1991). L'étude a démontré qu'un nombre

peu probable de travailleurs avaient déclaré des réductions dans leur taux de rémunération au cours de l'année. Dans ce cas, il est probable que les effets de la lisière découlent d'une déclaration inexacte des salaires et ils peuvent, dans certains cas, refléter l'effet de la déclaration par personne interposée.

On sait que d'autres caractéristiques comme les heures habituelles sont mal déclarées et ces caractéristiques sont des candidates probables pour de forts effets de la lisière. Une étude (Lemaître, 1988) a montré que même lorsqu'il n'y avait qu'une semaine entre les déclarations, environ le tiers des répondants déclaraient des heures habituelles différentes pour la même semaine de référence. Le taux de divergence différait à peine, même quand le répondant déclarait lui-même ses heures lors des deux occasions. Dans ce cas, on pourrait choisir de supposer qu'il n'y a pas de changement dans les heures pour le même emploi et tout simplement confirmer si l'emploi était toujours à plein temps ou à temps partiel.

Le fait que la caractéristique étudiée fasse vraisemblablement ou non l'objet d'analyses de transition pourrait être un critère pour décider si l'on doit rappeler des renseignements sur une caractéristique particulière. Il s'agit d'un point critique, parce que Hill et Hill (1986) ont trouvé que dans un modèle de hasards proportionnels des transitions de personnes qui trouvent un emploi, qui utilise des données de la SIPP, c'est le fait que la semaine de transition soit ou non une semaine de lisière qui était la variable prédictive la plus importante d'une transition. Malheureusement, il faut des pouvoirs de prescience considérables pour déterminer les variables qui ne seront jamais utilisées dans les analyses de transition.

Pour certaines caractéristiques, on pourrait signaler les changements à l'intervieweur (ce qui est possible dans le cadre de l'IPAO) et les faire confirmer par le répondant. Toutefois, cette stratégie peut causer des problèmes à l'intervieweur si de nombreux changements sont enregistrés et que le répondant

doit confirmer, à maintes reprises, un changement dont il saura qu'il ne peut être relevé que si l'intervieweur se reporte à ce qui a été déclaré lors de l'interview précédente. Il n'est pas vraisemblable que le répondant apprécierait, si de nombreuses confirmations de ce genre sont demandées, le message sous-entendu, peut-être involontaire, que les réponses qu'il fournit laissent à désirer.

7.2 La rétro-information comme partie intégrante de l'interview

En bref donc, il ne semble pas y avoir de réponse simple à savoir combien de renseignements rappeler aux répondants. Une chose est certaine, si l'on ne rappelle pas de renseignements, il se produira, assurément, des effets de la lisière. Toutefois, selon la façon de procéder, le rappel des renseignements n'a pas à être trop indiscret ou à comporter un fardeau de réponse excessif. L'IPAO offre la possibilité d'adapter les questions à des situations particulières. Par exemple, il n'est pas nécessaire de demander au répondant de confirmer ce qu'il a déclaré une année auparavant (comme cela a été fait pour les emplois dans l'EA), il suffit de supposer que la déclaration précédente était exacte et de réaliser l'interview à partir de ce point. Dans le cas des emplois, par exemple, on pourrait demander «Travaillez-vous toujours pour...?». Cette question se compare à ce qui est fait actuellement dans l'enquête sur la population active pour l'activité économique et la profession. Cette question pourrait être suivie d'une autre «Avez-vous été promu chez cet employeur?». Si la réponse à la première de ces questions est «Oui» et à la deuxième «Non», il y a beaucoup de questions qui n'ont pas à être posées à nouveau ou même pour lesquelles on n'a pas à confirmer les réponses, comme l'activité économique, la profession, la syndicalisation, l'adhésion à un régime de retraite, etc.

De plus, le processus de rétro-information ne devrait pas être considéré comme un mal nécessaire auquel on doit avoir recours seulement pour éviter les effets de la

lisière, mais comme une partie intégrante du processus d'interview. Une façon de visualiser une interview longitudinale consiste à considérer cette dernière comme un véhicule pour mesurer le changement. Dans cette perspective, l'interview repère établit essentiellement la situation de l'enquêté au moment de l'interview. Le rôle des interviews réalisées ultérieurement est alors de déterminer les changements qui se sont produits par la suite. La rétro-information constitue alors un outil opérationnel visant à aider le répondant à déclarer le changement avec exactitude et, en particulier, à assurer que le changement déclaré à la lisière entre deux interviews semble comparable au changement déclaré pendant la période visée par une seule interview.

8. Conclusion

La déclaration d'un changement excessif à la limite entre deux interviews est un problème auquel aucune enquête longitudinale réalisée jusqu'ici n'a échappé. Maintenant, avec un certain recul et avec nos connaissances actuelles sur les erreurs de réponse, il semble rétrospectivement que les personnes qui ont réalisé des enquêtes aient été d'une naïveté désespérante en s'attendant à ce que des données particulières soient «régulières» quand elles sont recueillies à deux occasions, à une année d'intervalle. Le mystère est de comprendre pourquoi le problème n'a pas été reconnu plus tôt, particulièrement puisque certaines enquêtes longitudinales (p. ex., la PSID et les National Longitudinal Surveys des États-Unis) sont réalisées depuis plus de vingt ans. La réponse est probablement qu'avec la rareté des données sur les sujets sociaux au moment où de telles enquêtes ont été mises en application, l'attention était beaucoup plus concentrée sur les résultats des mesures que sur le processus qui permettait d'obtenir ces dernières.

L'accent mis sur le problème de la lisière au cours des dernières années a attiré l'attention sur le fait que les enquêtes longitudinales doivent l'être non seulement au

sens de suivre les mêmes répondants dans le temps, mais aussi au sens de recueillir des données qui sont cohérentes dans le temps. En pratique, il semblerait que la seule façon fiable d'atteindre ce but soit de rappeler aux enquêtés des renseignements que ces derniers ont déjà fournis. On a toujours craint que cette façon de procéder n'entraîne une réduction des estimations du changement; cependant, les conséquences de ne pas appliquer une telle méthode sembleraient encore plus graves.

L'enquête sur l'activité a été un pionnier remarquable dans le domaine de l'interview avec rétro-information. En fait, la technique utilisée dans le cadre de cette enquête pour rappeler des renseignements sur les emplois constitue le seul cas documenté d'atténuation, en fait d'élimination, réussie des effets de la lisière pour une caractéristique particulière. Toutefois, puisque les effets de la lisière semblent être très répandus et qu'ils semblent associés à la majorité des caractéristiques, on doit disposer d'une stratégie plus générale. Il semble évident que le succès dans ce domaine exigera l'utilisation d'une méthode d'interview avec rétro-information de type général comme partie intégrante du processus d'interview.

BIBLIOGRAPHIE

Bootby, Daniel (1991). Communication personnelle.

Burkhead, D. and J. Coder (1985). Gross Changes in Income Recipiency from the Survey of Income and Program Participation. Proceedings of the Social Statistics Section, American Statistical Association, 351-356.

Czajka, J. (1983). Subannual Income Estimation, Peper 7 in M. David (ed.): Technical, Conceptual and Administrative Lessons of the Income Survey Development Program (ISDP). New York: Social Science Research Council.

Goudreau, J., H. Oberheu and D. Vaughan (1984). An Assessment of the Quality of Survey Reports of Income from the Aid to Families with Dependent Children (AFDC) Program. Journal of Business and Economic Statistics, Vol. 2, No. 2, 179-186.

Hill, D. (1987). Response Errors Around the Seam: Analysis of Change in a Panel with Overlapping Reference Periods. Proceedings of the Section on Survey Research Methods, American Statistical Association, 210-215.

Hill, M.S. and D.H. Hill (1986). Labour Force Transitions: A Comparison of Estimates from Two Longitudinal Surveys. Proceedings of the Section on Survey Research Methods, American Statistical Association, 220-225.

Kalton, Graham and Michael E. Miller (1991). The Seam Effect with Social Security Income in the Survey of Income and Program Participation. Journal of Official Statistics, Vol.7 No.2, 235-245.

Lemaître, Georges (1988). A look at response errors in the Labour Response Survey. The Canadian Journal of Statistics, Vol.16, Supplement 127-141.

Marquis, K. (1992). Communication personnelle.

Marquis, K. and J. Moore (1989). Response Errors in SIPP: Preliminary Results. Proceedings of the U.S. Bureau of the Census Annual Research Conference, 515-536.

Marquis, K. and J. Moore (1990). Measurement Errors in SIPP Program Reports. Proceedings of the U.S. Bureau of the Census Annual Research Conference, 721-745.

Marquis, Kent H., Jeffrey C. Moore and Vicki J. Huggins (1990). Implications of SIPP Record Check Results for Measurement Principles and Practice. Proceedings of the Section on Survey Research Methods, American Statistical Association, 564-569.

Moore, J. (1990). An Examination of Proxy Effects in Reporting of Income Recipiency (Exact title?). Paper presented at International Conference on Measurement Errors in Surveys, Tucson, Arizona, Nov. 1990).

Moore, J. and D. Kasprzyk (1984). Month-to-Month Recipiency Turnover in the ISDP. Proceedings of the Section on Survey Research Methods, American Statistical Association, 726-731.

Murray, T. Scott, Sylvie Michaud, Mike Egan and Georges Lemaître (1991). Invisible Seams? The Experience with the Canadian Labour Market Activity

Survey. Proceedings of the US Bureau of the Census Annual Research Conference, 715,730.

NCEUS (1979). Counting the Labor Force. National Commission on Employment and Unemployment Statistics. Washington. 214-217.

Neter J. and J. Waksberg (1965). Response errors in collection of expenditures data by household interviews: An experimental study. Bureau of the Census Technical Paper No. 11, U.S. Government Printing Office, Washington.

PCGF (1985). Proceedings of the Conference on Gross Flows in Labour Force Statistics. Bureau of the Census and Bureau of Labor Statistics.

Weidman, L. (1986). Investigation of Gross Changes in Income Recipiency from the Survey of Income and Program Participation. Proceedings of the Section on Survey Research Methods, American Statistical Association, 231-236.